Using under-exploited satellite observations to learn hybrid models and novel state variables: a toolbox for future physical-empirical forecasting systems

Alan Geer

Thanks to: Phil Browne, Tracy Scanlon, Nils Risse, Niels Bormann, Sebastien Massart, Tsz Yan Leung

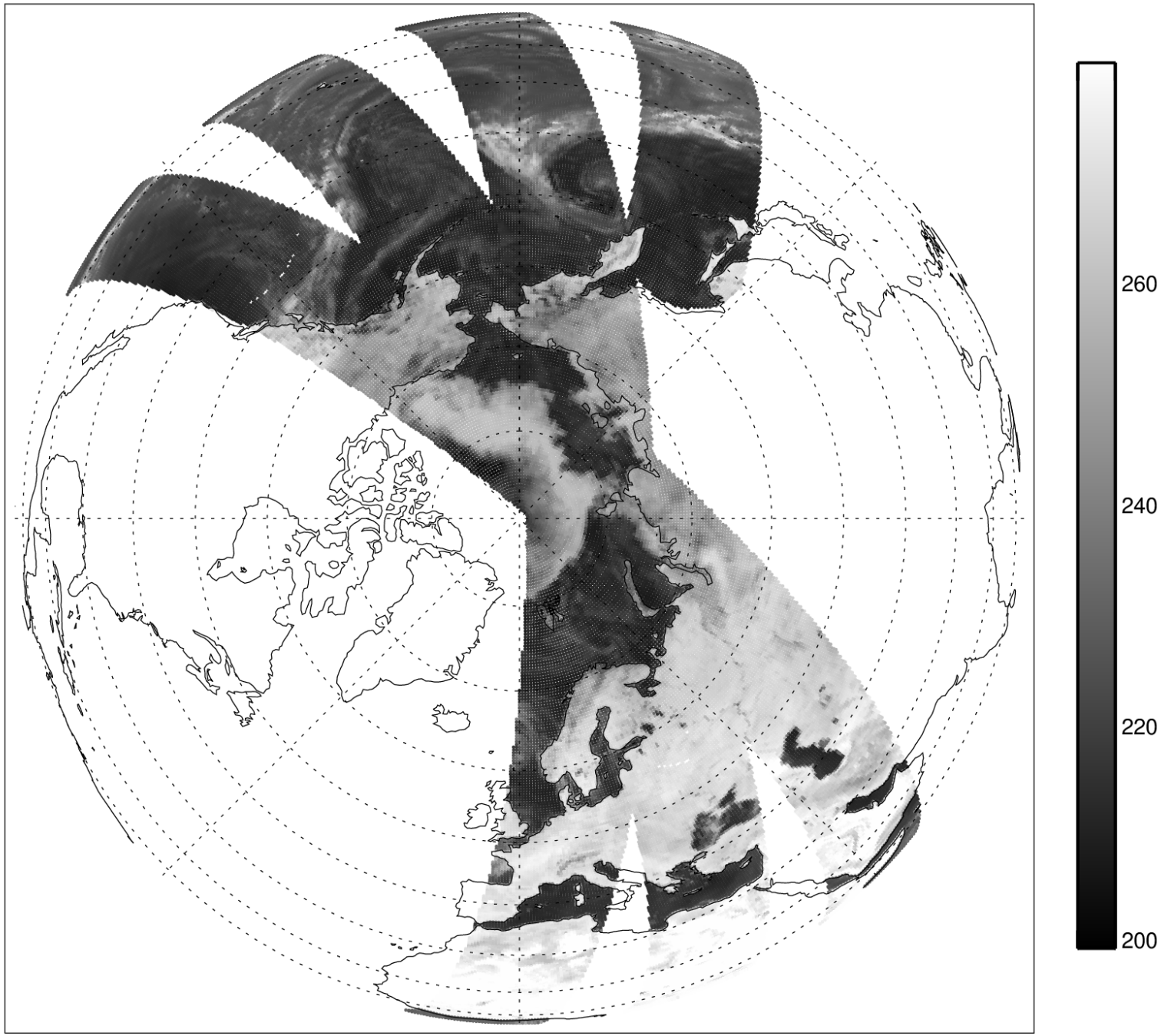
17 September 2026

ECMWF Annual Seminar 2026, Shinfield Park, Reading

Advanced Scanning Microwave Radiometer (AMSR2)

Observation composite for 1st Nov 2021

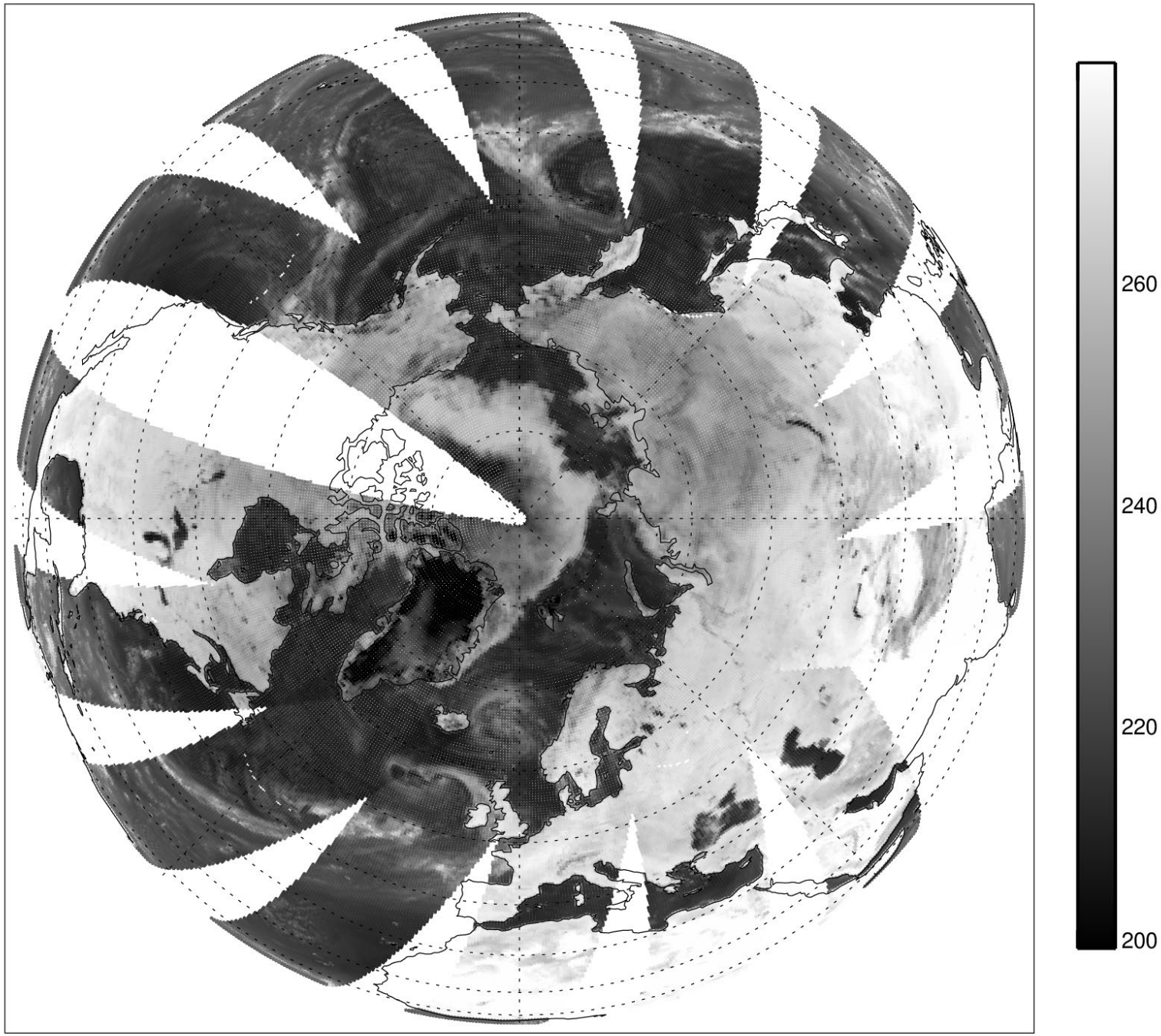
Brightness temperatures (TB) [Kelvin] at 37 GHz, v-polarised



Advanced Scanning Microwave Radiometer (AMSR2)

Observation composite for 1st Nov 2021

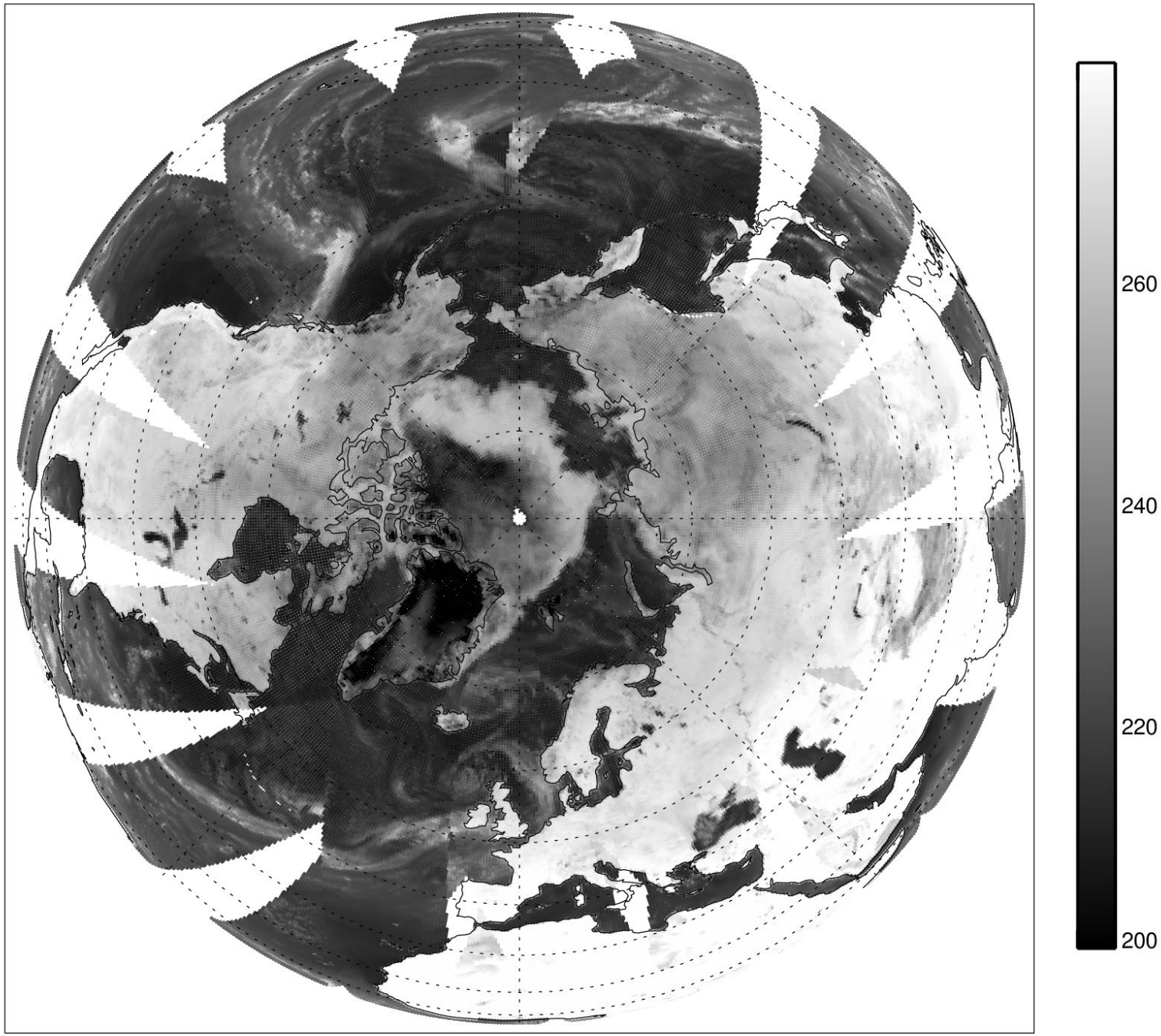
Brightness temperatures (TB) [Kelvin] at 37 GHz, v-polarised



Advanced Scanning Microwave Radiometer (AMSR2)

Observation composite for 1st Nov 2021

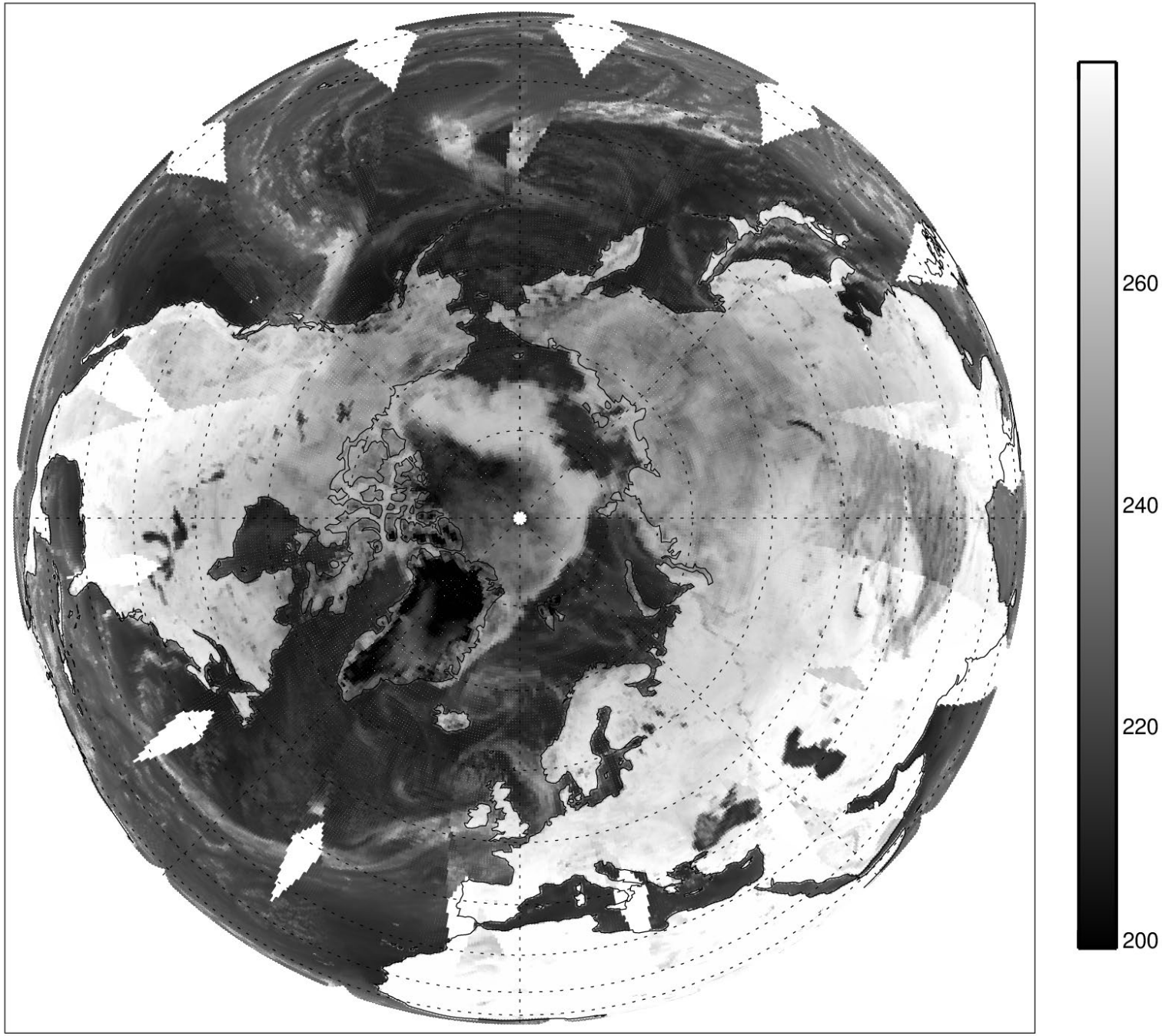
Brightness temperatures (TB) [Kelvin] at 37 GHz, v-polarised



Advanced Scanning Microwave Radiometer (AMSR2)

Observation composite for 1st Nov 2021

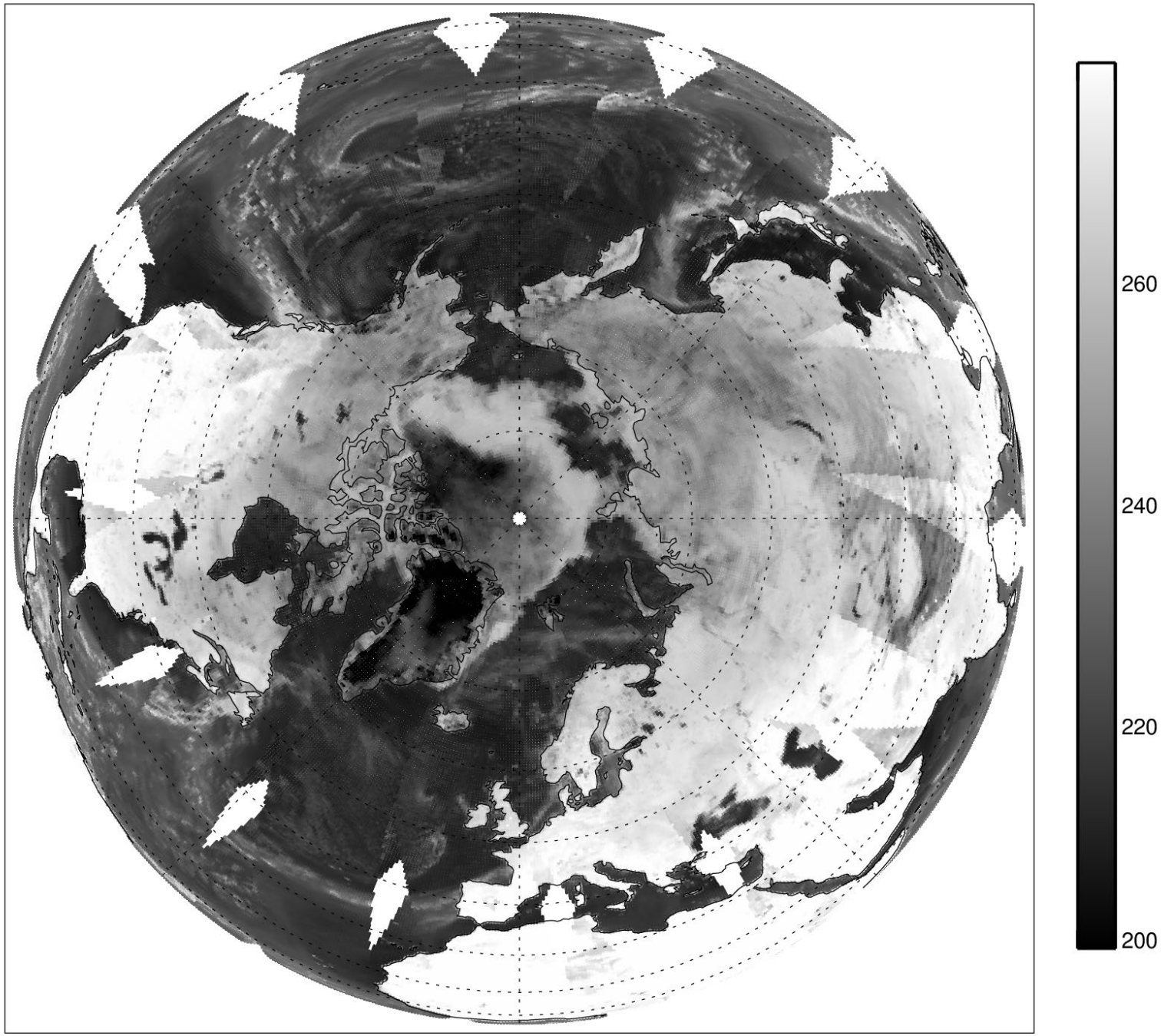
Brightness temperatures (TB) [Kelvin] at 37 GHz, v-polarised



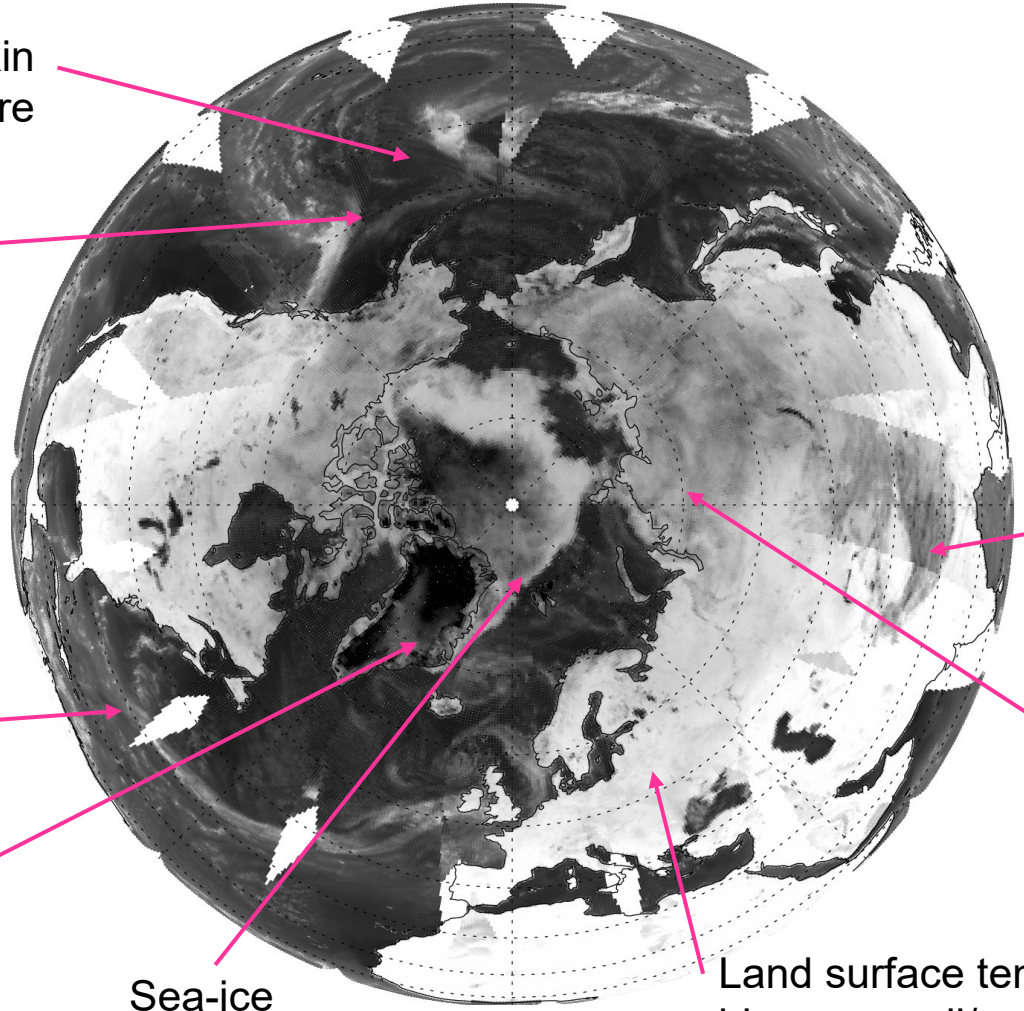
Advanced Scanning Microwave Radiometer (AMSR2)

Observation composite for 2nd Nov 2021

Brightness temperatures (TB) [Kelvin] at 37 GHz, v-polarised

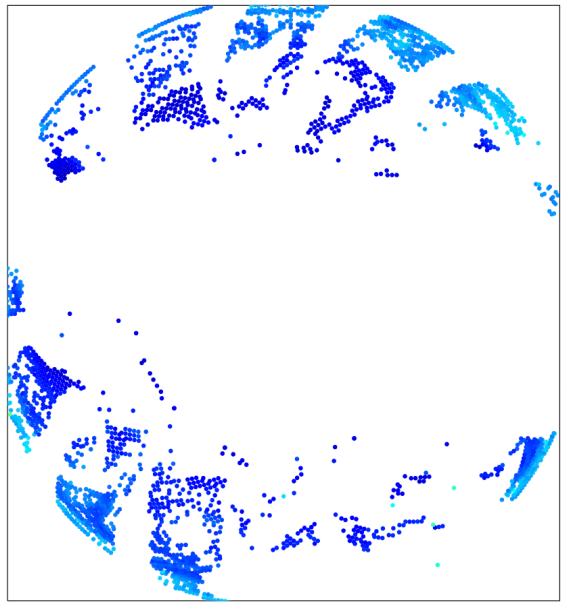
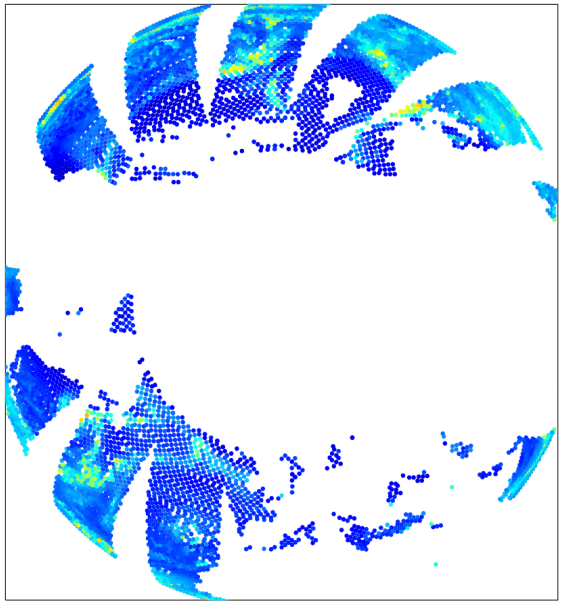
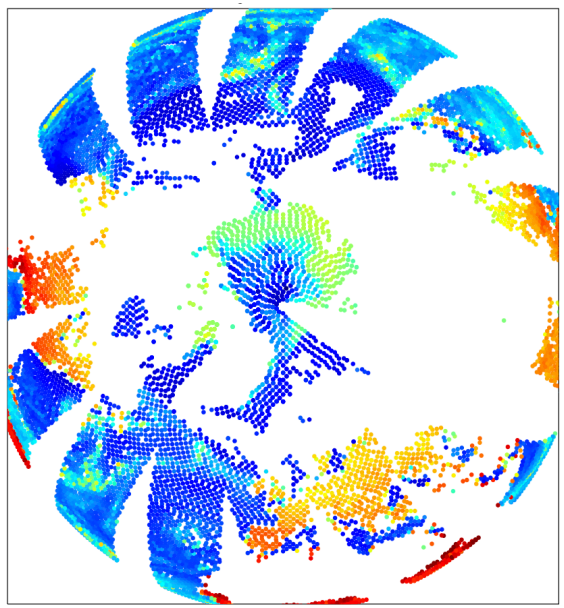
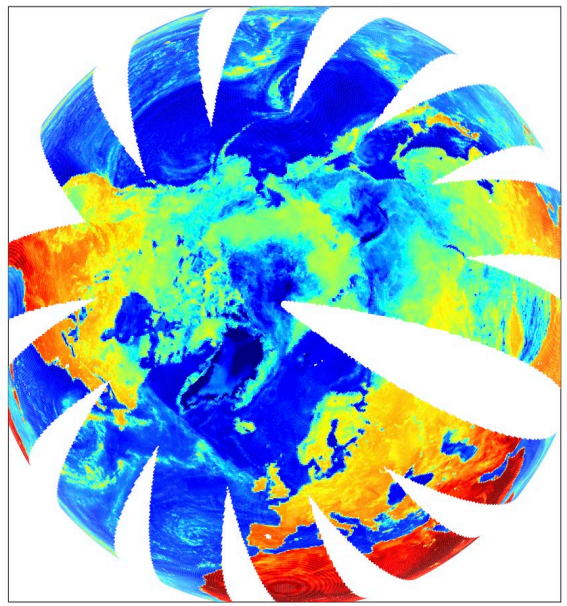


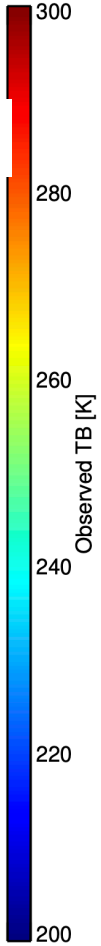
AMSR2 observation composite for 2nd Nov 2021
37 GHz, v-polarised



All-sky all-surface assimilation at ECMWF

Add coverage

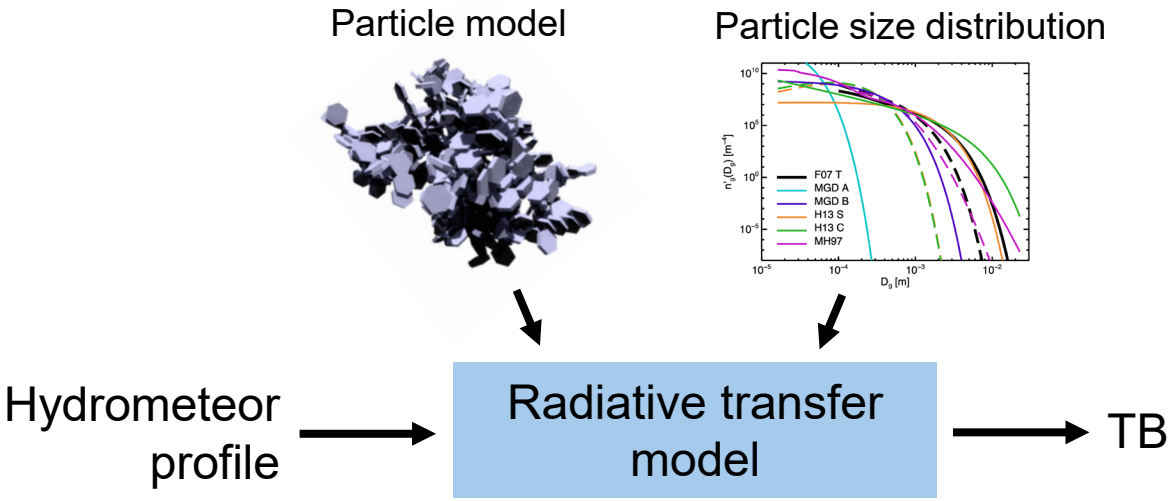
Clear-sky ocean	All-sky ocean	Cycle 50r1	All-sky all-surface all-obs
			
1990s	2009	2026	ongoing
water vapour temperature surface wind speed	+ cloud precipitation	+ sea ice concentration ocean skin temperature	+ everything else



Add data assimilation sensitivity

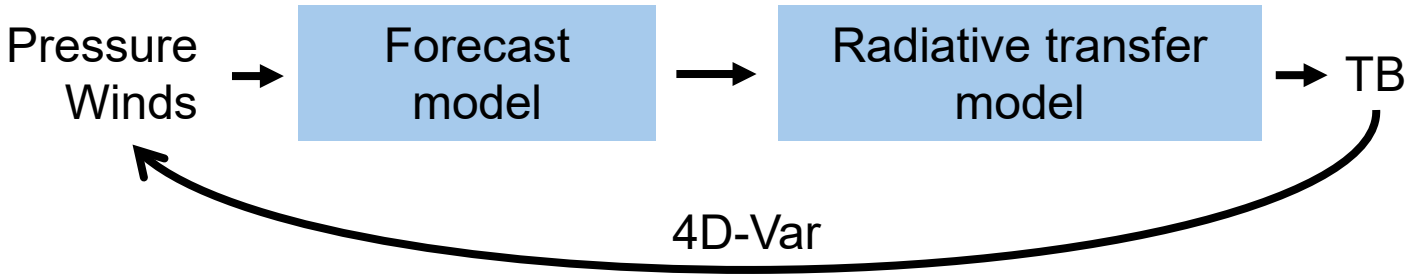
Lessons from all-sky assimilation

- Forward models depend on poorly-known, situation-dependent, micro-scale parameters:
  - Some form of parameter estimation is needed



- Don’t be too afraid of highly nonlinear, non-Gaussian, non-predictable problems
  - Incremental 4D-Var, simplified linearized physics, observation error inflation…

- 4D-Var provides model tracing:
  - 4D-Var is a sophisticated Bayesian retrieval system



ARTS large plate aggregate: https://doi.org/10.5194/essd-10-1301-2018

All-surface assimilation: sea ice

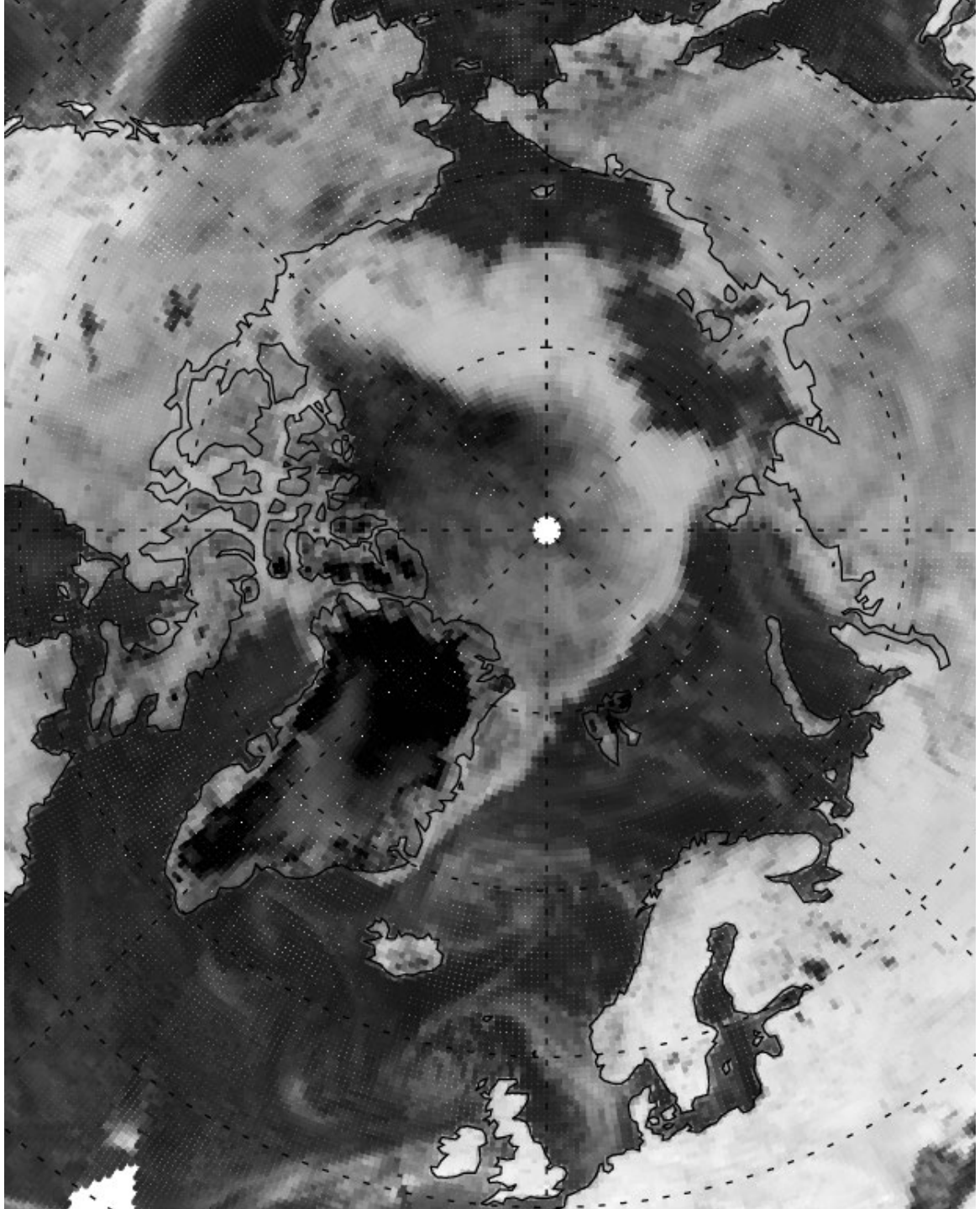
- In absence of any atmosphere:

$$\mathrm{TB} = C\epsilon_{\mathrm{ice}}T_{\mathrm{sfc}} + (1-C)\epsilon_{\mathrm{ocean}}T_{\mathrm{sfc}}$$

Sea ice concentration

Sea ice surface emissivity

- For the purposes of this work, we assume surface temperature T_{sfc} is known (e.g. from IFS short forecast)



Full modelling of the surface emission?

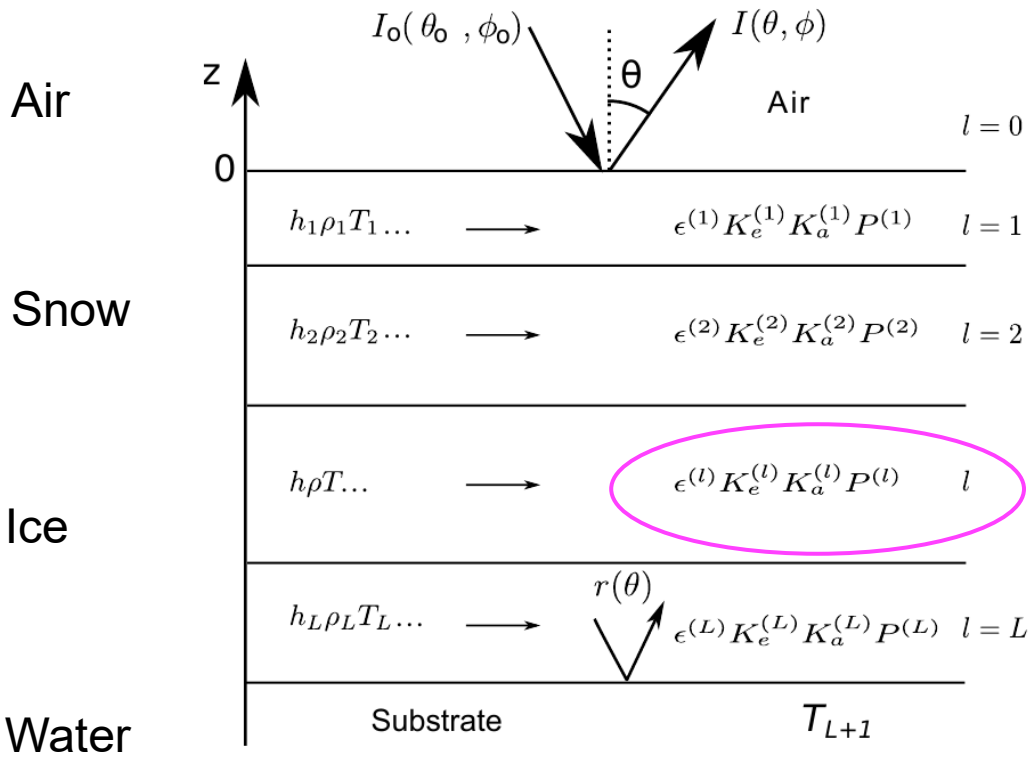


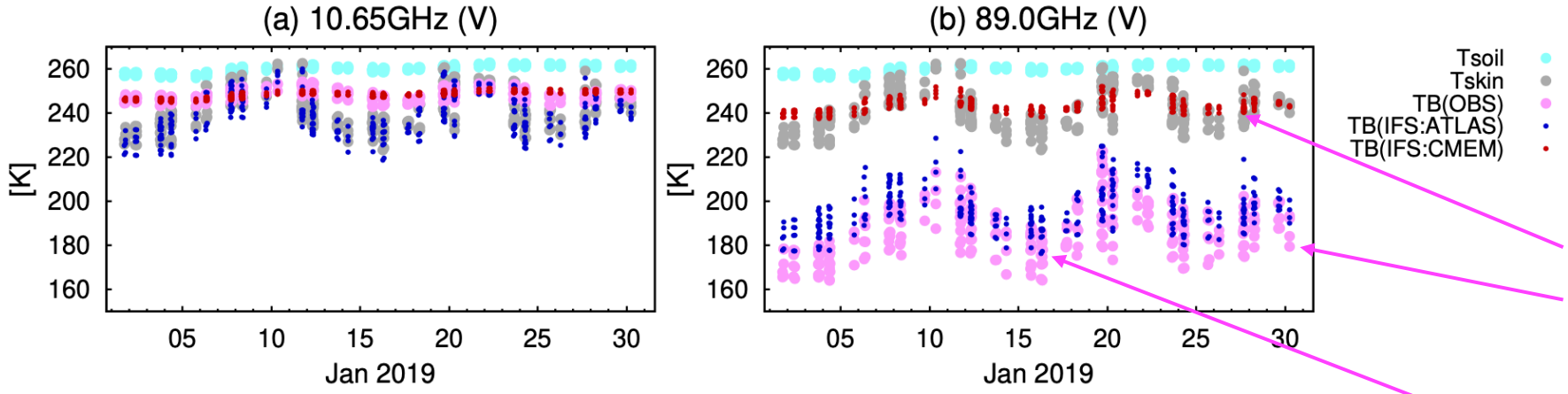
Figure 1. Multilayered medium modeled by SMRT. The incident radiation I_o comes either from a radar beam (active mode) or from the sky (passive mode with atmospheric contribution).

Optical properties: extinction, absorption, phase function

But where do we get profiles of snowpack and sea ice optical properties?

Snow Microwave Radiative Transfer model
Picard et al. (2018, https://doi.org/10.5194/gmd-11-2763-2018)

Global scale physical snow modelling is still in development



At an Arctic location (2x2 degrees box): timeseries of modelled and observed brightness temperature

Hirahara et al. (2020, https://doi.org/10.3390/rs12182946)

CMEM physical model and observations do not agree

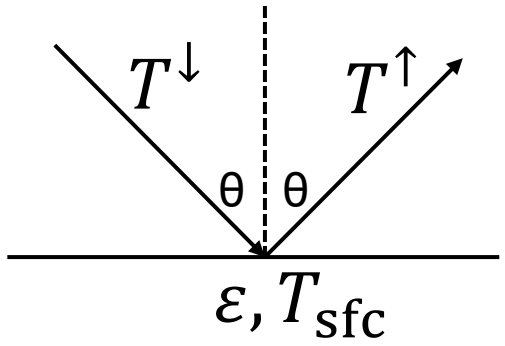
Emissivity atlas does best in fitting the observations

Empirical emissivity: retrieval

- Retrieval → surface emissivity atlas
- Dynamic emissivity retrieval in a window channel → dynamic surface emissivity in sounding channel
  - Not useful when we need to assimilate those window channels
- Other issues with emissivity retrieval:
  - What are the retrieval errors and how do they vary with the situation?
  - If the observations do not see the surface at all → unphysical retrievals

$$\varepsilon = \frac{(T^{\uparrow} - T^{\downarrow})}{(T_{\mathrm{sfc}} - T^{\downarrow})}$$

TB upwelling $T^{\uparrow}$ and downwelling $T^{\downarrow}$ at the surface can be estimated from observations and radiative transfer modelling



$$T^{\uparrow} = (1 - \varepsilon)T^{\downarrow} + \varepsilon T_{\mathrm{sfc}}$$

The need for a constrained emissivity spectrum

$$\begin{pmatrix} \mathrm{TB}_1 \\ \mathrm{TB}_2 \end{pmatrix} = CT_{\mathrm{sfc}} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

An even simpler illustrative sea ice concentration retrieval, from two channels, assuming the atmosphere is optically thin and the ocean is purely reflective

Two observations (TB_1 and TB_2) and two unknowns (ϵ_1 and ϵ_2)
→ no information left to retrieve sea ice concentration

But if emissivity has a known spectral dependence, e.g.

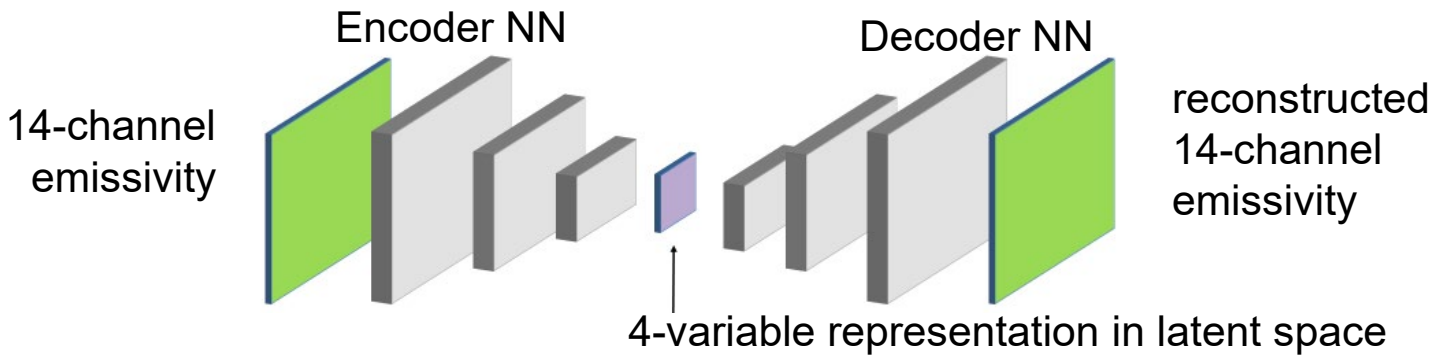
$$\epsilon_1 = \mathrm{f}(\epsilon_2)$$

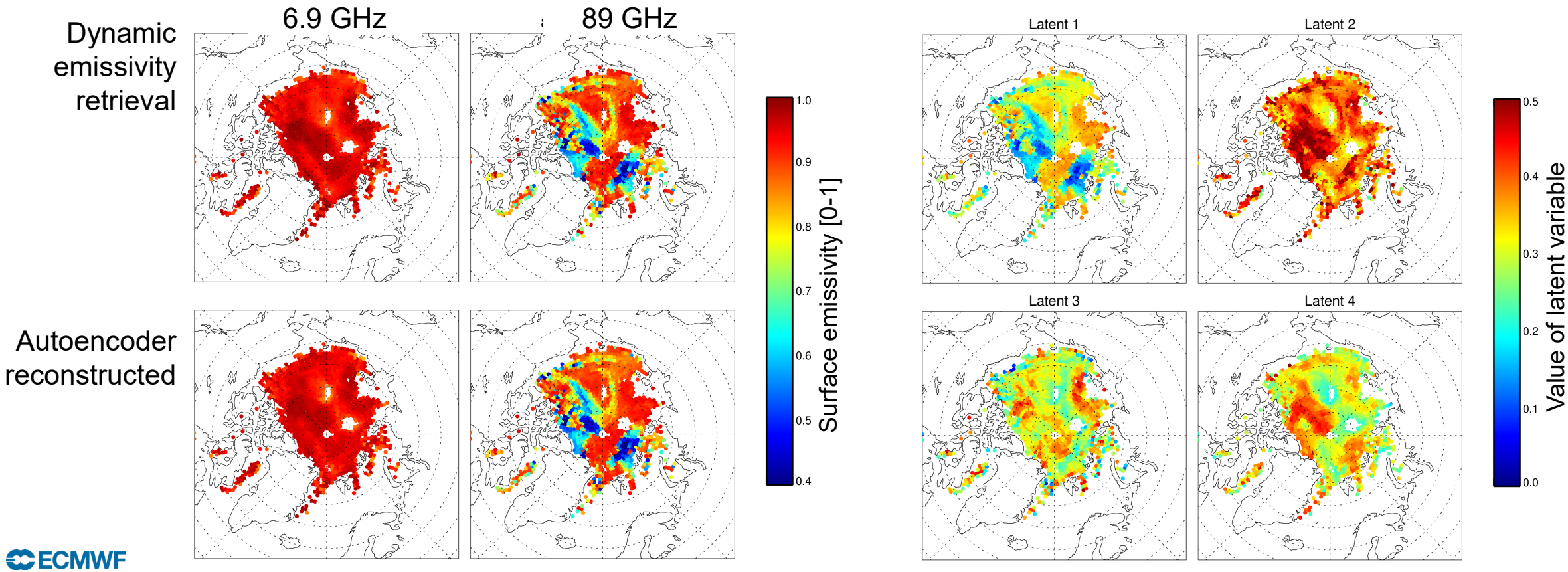
then we can infer sea ice (1 unknown) and emissivity (1 unknown) together

We could use principal components of measured emissivity spectra

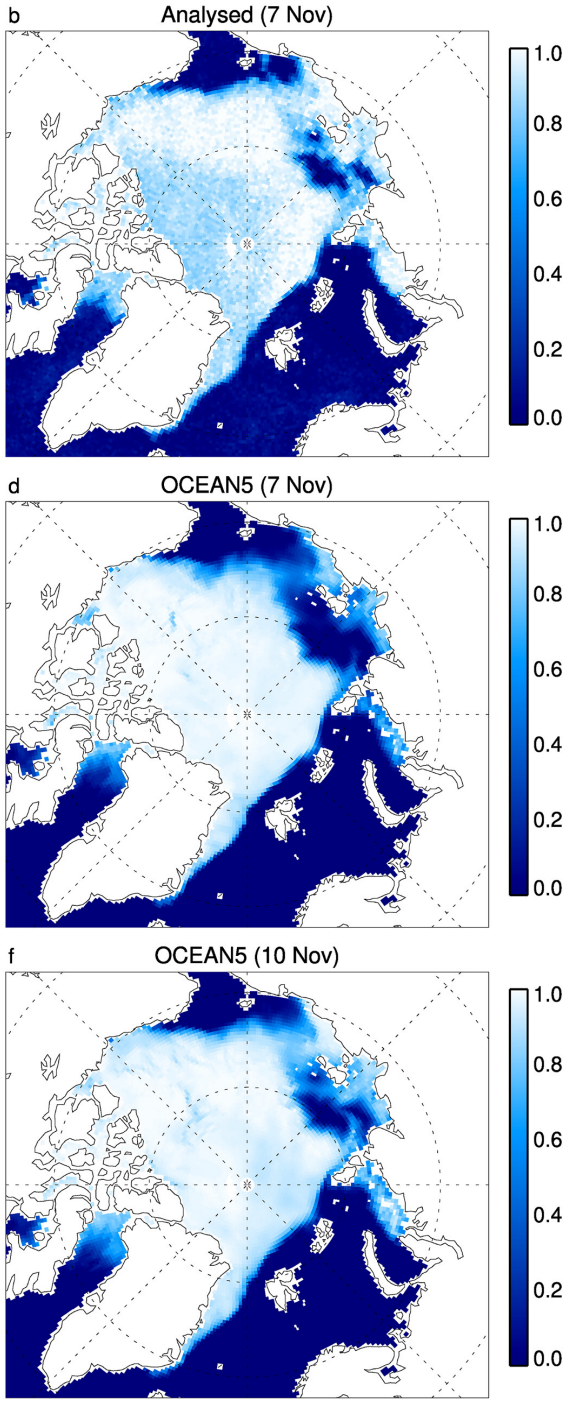
- E.g. in infrared DA: Pavelin and Candy (2014) https://doi.org/10.1002/qj.2218
- PCs give a **linear** reconstruction of the emissivity

Nonlinear constrained spectral dependence: autoencoder





ECMWF



Problems with the autoencoder

- Emissivity is a retrieval → need situation-dependent errors
- How to handle sensors observing different frequencies, times, and locations? SMAP, AMSR2, GMI etc.
- We need specifically sea ice emissivity, not field of view emissivity
- So we need sea ice concentration to separate sea ice and ocean emissivity
  - But sea ice concentration datasets have historically been unreliable
  - E.g. OCEAN5 had a lag of 48 h – 72 h

→ Represent observation errors

→ Learn directly from TB

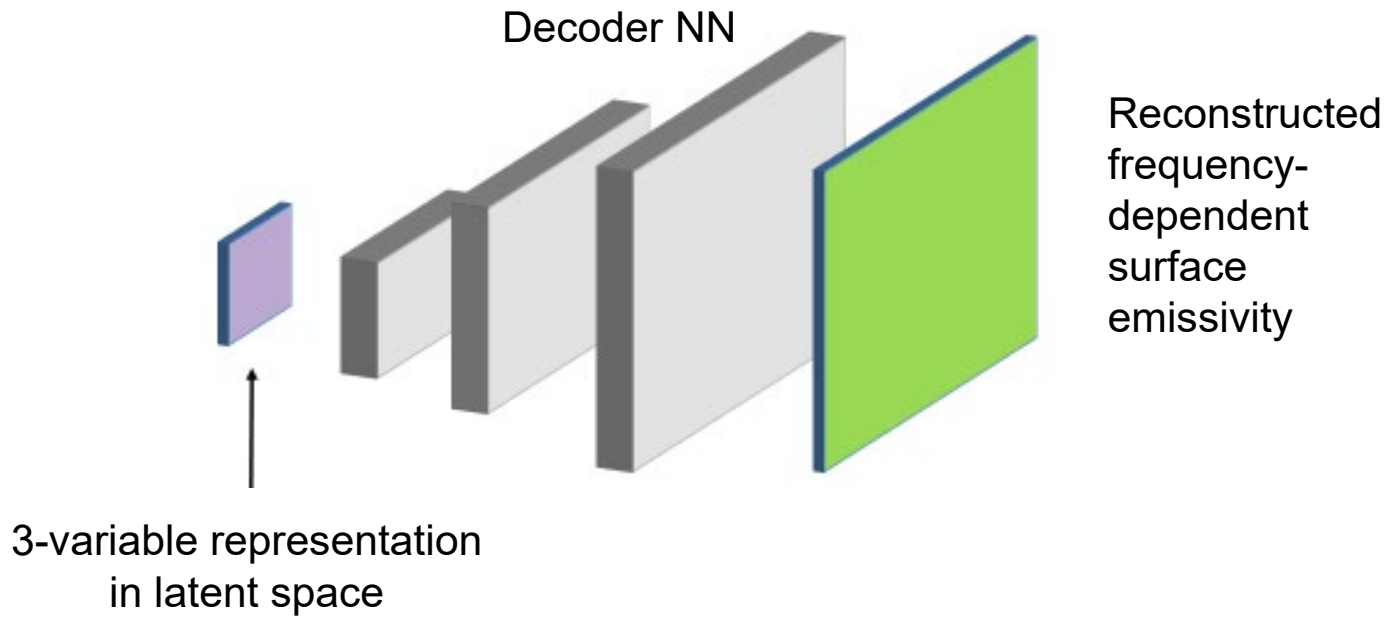
→ Use an observation operator to represent interpolation and radiative transfer

→ Infer sea ice emissivity on a grid

→ Infer sea ice concentration from the observations at the same time

We need data assimilation!

But how do we constrain the spectral dependence of surface emissivity? We also need machine learning



Let’s put the latent space on a space and time grid too → **Empirical state**

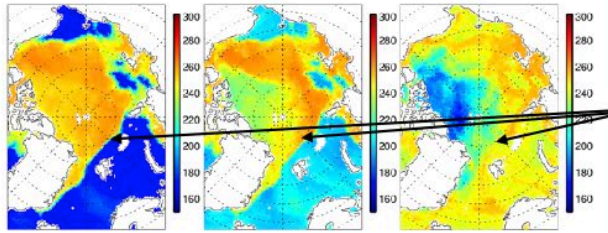
One way to combine machine learning and DA in practice

- Python framework using Keras and Tensorflow
  - Implement both physical and empirical components as network layers
- Minimisation method:
  - Stochastic gradient descent (Adam)
  - Free from Gaussian and linear assumptions
- Objective
  - Jo loss term + targeted prior constraints (~B ~regularization)
- Gradient computation (e.g. TL/AD/back-propagation):
  - Auto-differentiation using the Tensorflow framework
- Computational efficiency
  - Tensorflow backend runs optimised and highly parallel (~vector parallelism) on GPU

Physical (Bayesian) network representation of sea ice and snow radiative transfer

AMSR2 observations

$$J_{\text{obs}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(y_{\text{obs},ij} - y_{\text{sim},ij})^2}{r_j^2}$$



Known atmosphere radiative transfer

$\mathbf{y} = f_{\text{atmos}}(\mathbf{e}, T_S, \mathbf{z}_A)$

Known atmosphere

Mixed surface emissivity

$\mathbf{e} = C_{\text{ice}}\mathbf{e}_{\text{ice}} + (1 - C_{\text{ice}})\mathbf{e}_{\text{wat}}$

Known water surface emissivity

Sea ice concentration at observation location

Sea ice surface emissivity

Fixed parameter

Dependent parameter

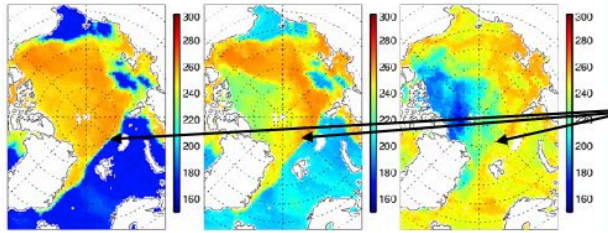
Trainable parameter

Geer (2024, https://doi.org/10.1029/2023MS004080)

https://github.com/ecmwf-projects/empirical-state-learning-seaice-emissivity-model

Physical (Bayesian) network representation of sea ice and snow radiative transfer

AMSR2 observations



$$J_{\text{obs}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(y_{\text{obs},ij} - y_{\text{sim},ij})^2}{r_j^2}$$

Known atmosphere radiative transfer

$\mathbf{y} = f_{\text{atmos}}(\mathbf{e}, T_S, \mathbf{z}_A)$

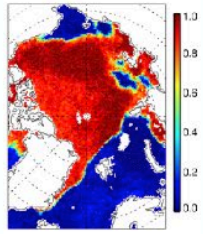
T_S $\mathbf{z}_A$ Known atmosphere

Mixed surface emissivity

$\mathbf{e} = C_{\text{ice}}\mathbf{e}_{\text{ice}} + (1 - C_{\text{ice}})\mathbf{e}_{\text{wat}}$

$\mathbf{e}_{\text{wat}}$ Known water surface emissivity

Map of sea ice fraction to be estimated



$\mathbf{C}_{\text{ice}}$ C_{ice} $C_{\text{ice}} = h(\mathbf{C}_{\text{ice}}, \mathbf{z}_G)$ $\mathbf{e}_{\text{ice}}$ Sea ice surface emissivity

$\mathbf{z}_G$ Geolocation

h() Observation operator: map to observation location in time and space

Fixed parameter

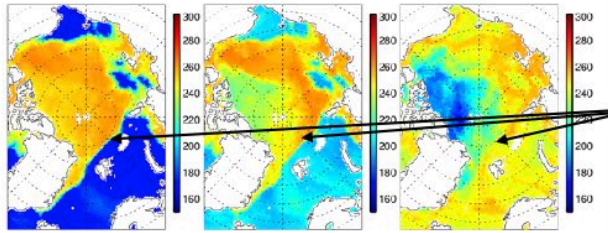
Dependent parameter

Trainable parameter

The whole trainable empirical-physical network

AMSR2 observations

$$J_{\text{obs}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(y_{\text{obs},ij} - y_{\text{sim},ij})^2}{r_j^2}$$



Known atmosphere radiative transfer

$\mathbf{y} = f_{\text{atmos}}(\mathbf{e}, T_S, \mathbf{z}_A)$

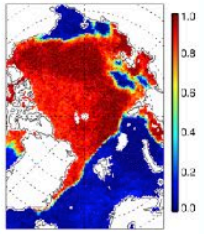
Known atmosphere

Mixed surface emissivity

$\mathbf{e} = C_{\text{ice}}\mathbf{e}_{\text{ice}} + (1 - C_{\text{ice}})\mathbf{e}_{\text{wat}}$

Known water surface emissivity

Map of sea ice fraction to be estimated

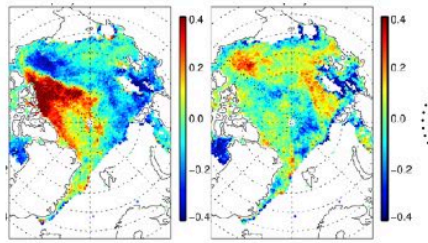


$C_{\text{ice}} = h(\mathbf{C}_{\text{ice}}, \mathbf{z}_G)$

$\mathbf{e}_{\text{ice}} = f_{\text{empirical}}(\mathbf{w}, \mathbf{x}_{\text{ice}}, \mathbf{z}_B)$

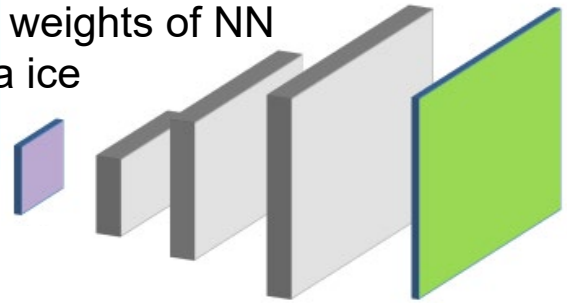
Geolocation

Maps of empirical state parameters representing **unknown** sea ice state including microstructure



$\mathbf{x}_{\text{ice}} = h(\mathbf{X}_{\text{ice}}, \mathbf{z}_G)$

w – trainable weights of NN model for sea ice



h() Observation operator: map to observation location in time and space

Fixed parameter

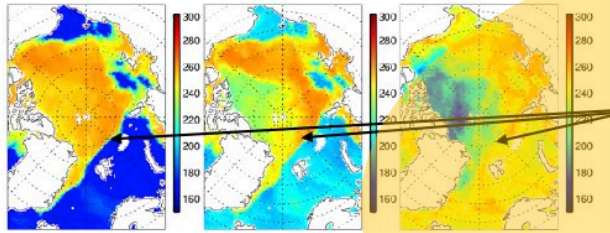
Dependent parameter

Trainable parameter

A tightly coupled physical / machine learning hybrid model in a data assimilation framework, implemented with ML

Physical model

AMSR2 observations



$$J_{\text{obs}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(y_{\text{obs},ij} - y_{\text{sim},ij})^2}{r_j^2}$$

Known atmosphere radiative transfer

$\mathbf{y} = f_{\text{atmos}}(\mathbf{e}, T_S, \mathbf{z}_A)$

T_S $\mathbf{z}_A$ Known atmosphere

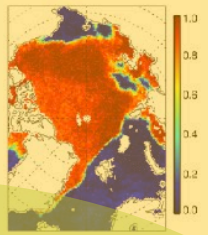
Mixed surface emissivity

$\mathbf{e} = C_{\text{ice}}\mathbf{e}_{\text{ice}} + (1 - C_{\text{ice}})\mathbf{e}_{\text{wat}}$

$\mathbf{e}_{\text{wat}}$ Known water surface emissivity

Physical space

Map of sea ice fraction to be estimated



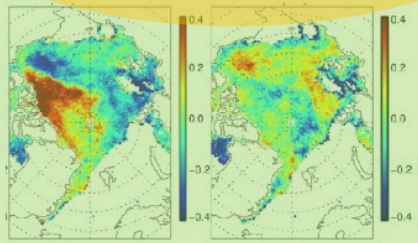
$\mathbf{C}_{\text{ice}}$ C_{ice} $C_{\text{ice}} = h(\mathbf{C}_{\text{ice}}, \mathbf{z}_G)$

$\mathbf{z}_G$ Geolocation

$\mathbf{e}_{\text{ice}}$ $\mathbf{e}_{\text{ice}} = f_{\text{empirical}}(\mathbf{w}, \mathbf{x}_{\text{ice}}, \mathbf{z}_B)$

$\mathbf{z}_B$

Maps of empirical state parameters representing **unknown** sea ice state including microstructure



$\mathbf{X}_{\text{ice}}$ $\mathbf{x}_{\text{ice}}$ $\mathbf{x}_{\text{ice}} = h(\mathbf{X}_{\text{ice}}, \mathbf{z}_G)$

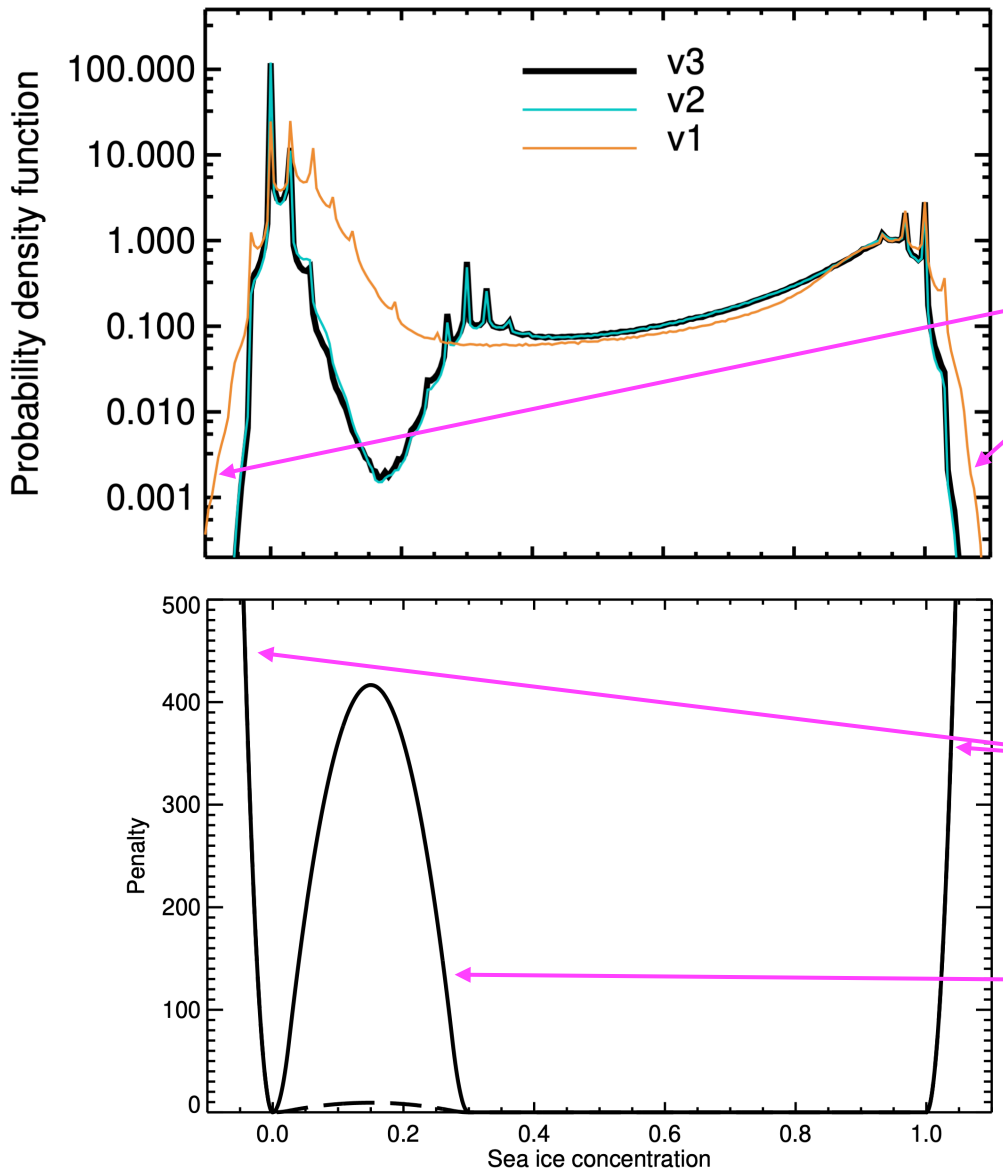
$\mathbf{w}$ w – trainable weights of NN model for sea ice

h() Observation operator: map to observation location in time and space

Empirical model

Latent space

Regularization loss terms = priors ≃ background error terms in DA



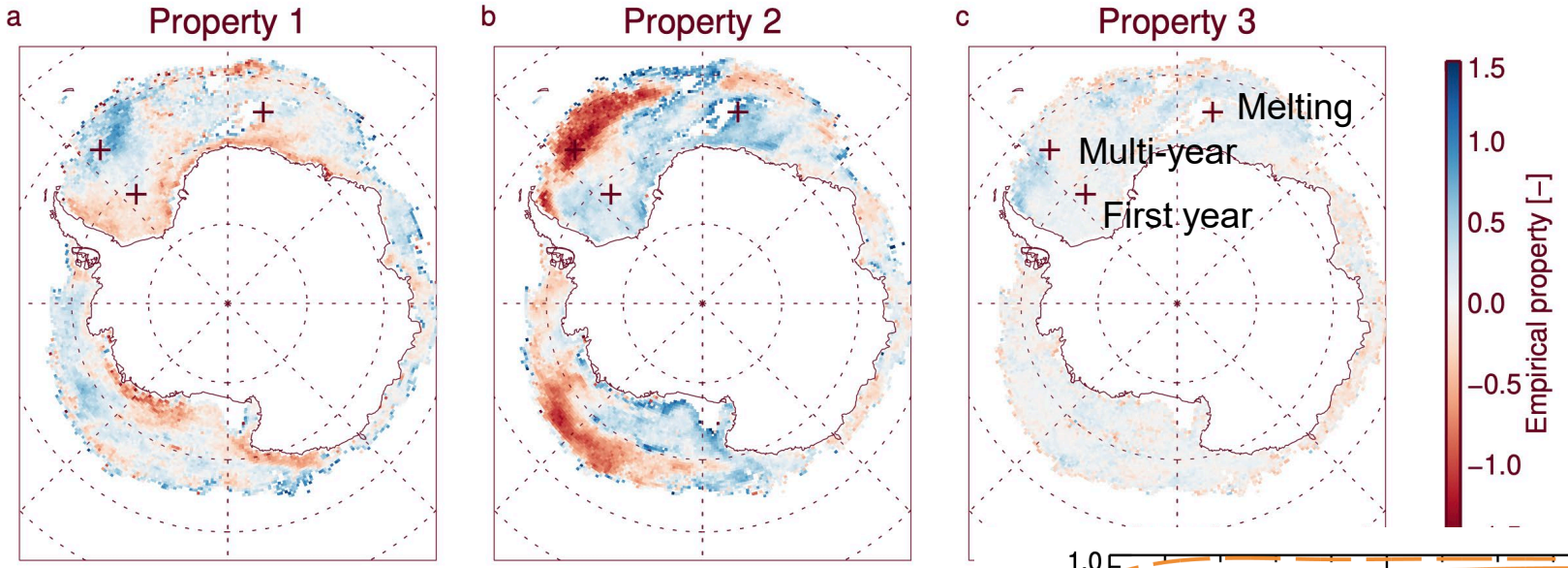
How to stop unphysical sea ice concentrations?

Bounds loss function (tighter in v2 and v3)

Bump function added in v2 and v3 to prevent spurious low-fraction sea ice

Emissivity and the empirical state

Example: 8th December 2020

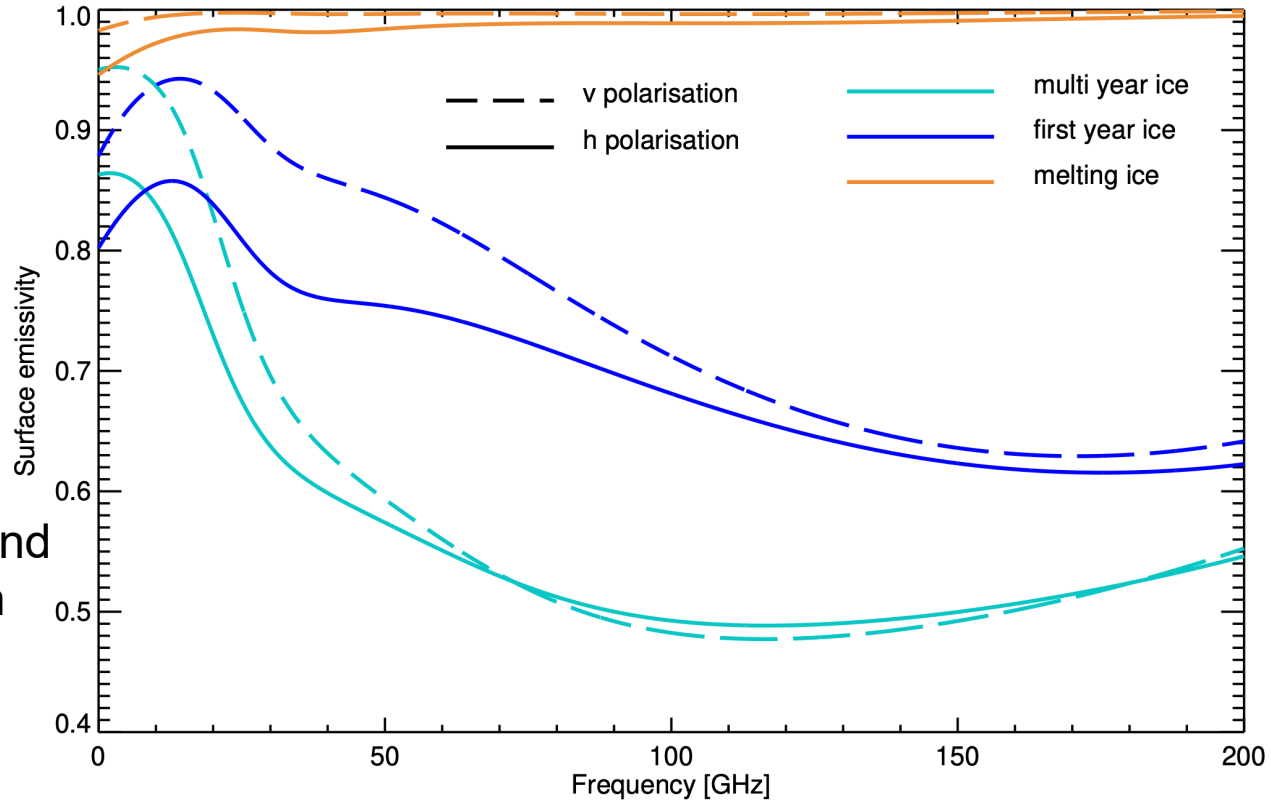


Three empirical properties



Skin temperature and frequency

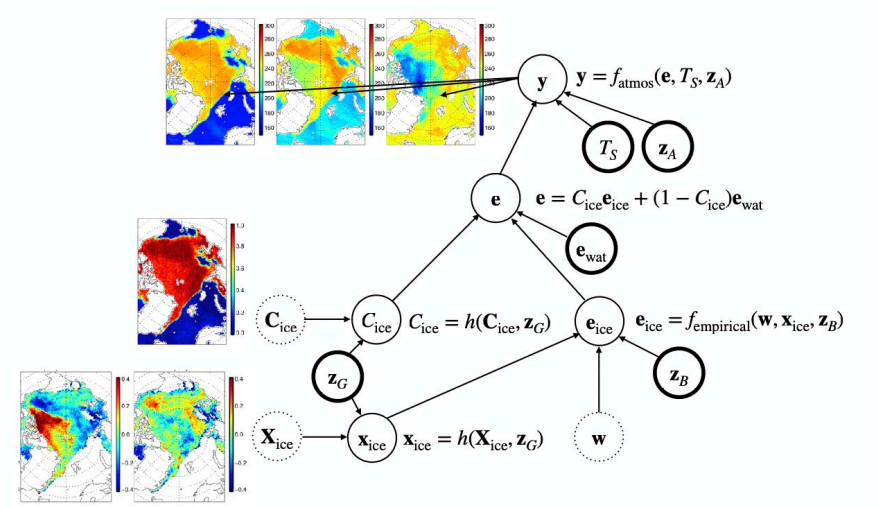
Surface emissivity spectrum and polarisation



The trainable parameters (for one year of training data)

Details of trainable parameters in the v3 configuration

Trainable parameter	Dimensions	Number of parameters
Sea ice concentration at grid locations	62,499 locations by 367 days	22,937,133
Empirical ice properties at grid locations	62,499 locations by 366 days by 2 properties	45,749,268
Empirical ice properties at observation locations	144,140,850 observation locations	144,140,850
Sea ice emissivity neural network	See Table 1	58
Ocean emissivity bias correction	23 channels	23
Sensor bias correction	23 channels by 4 sensors by 2 surface types	184
TOTAL		212,827,516

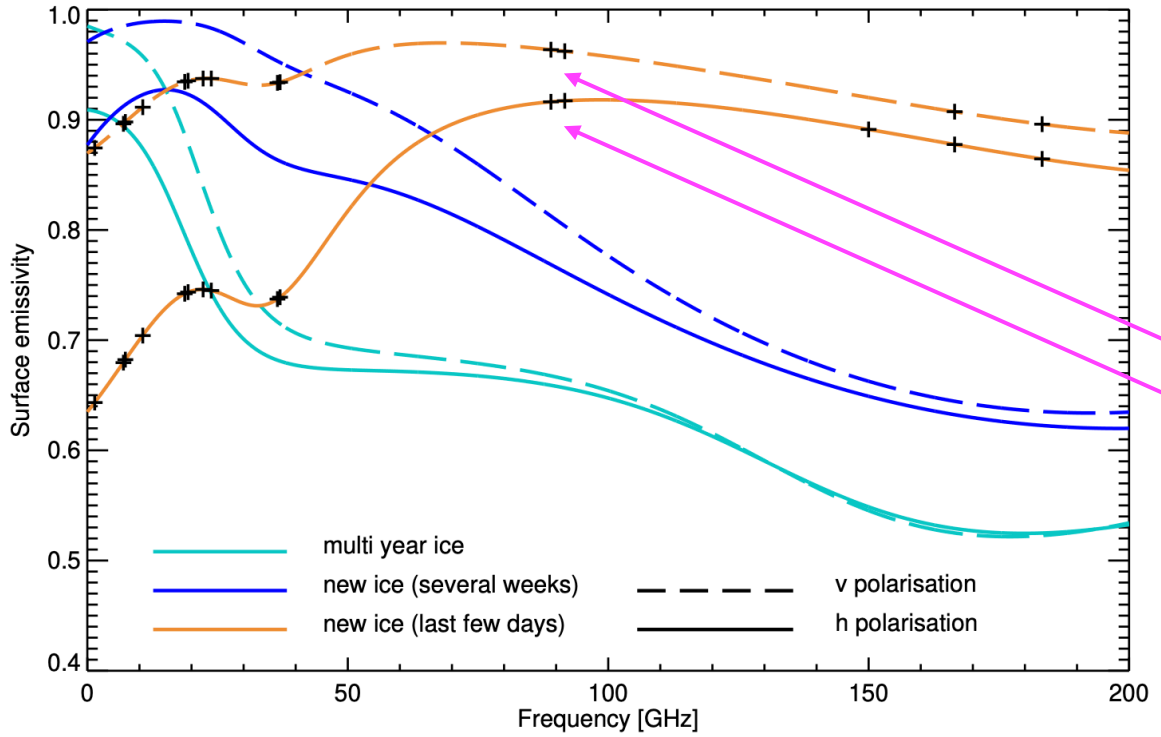


		parameters
Input vector	5 elements	–
Matrix multiplication	7 by 5 elements	35
Linear offset (bias)	7 elements	7
Sigmoid function (activation)	7 elements	–
Matrix multiplication	2 by 7 elements	14
Linear offset (bias)	2 elements	2
Sigmoid function (activation)	2 elements	–
Output vector	2 elements	–
TOTAL		58

Training data – observations

+ atmospheric terms, ocean surface emissivity and skin temperature at observation locations from IFS (ODB data)

Example emissivity spectra 8th December 2020



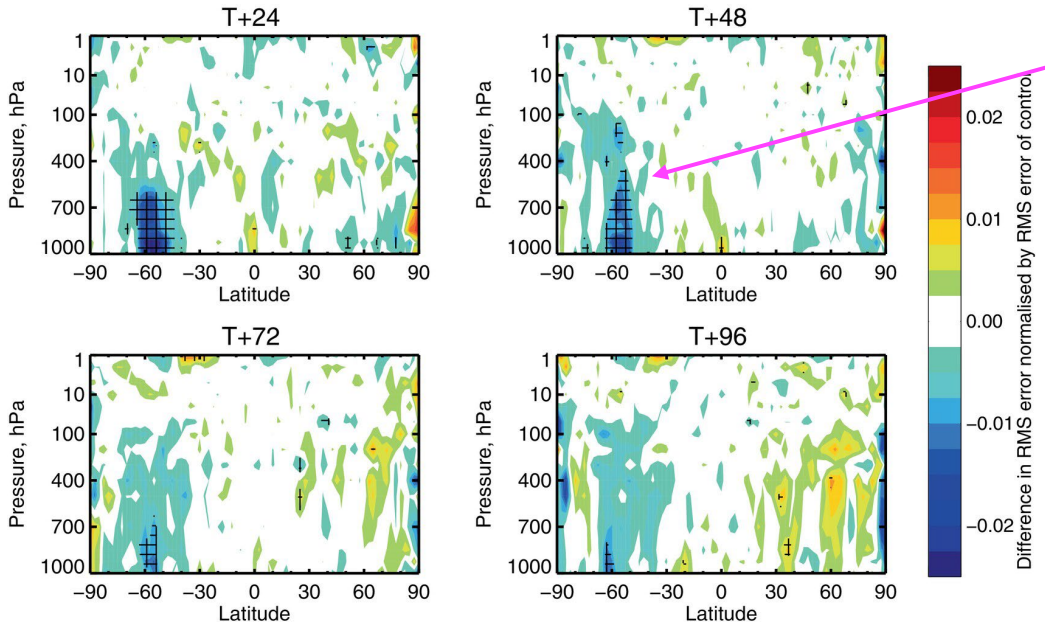
Channel label	Actual frequency on sensor [GHz] SMAP	AMSR2	GMI	SSMIS	Observation error [K]
1.4v	1.4135	–	–	–	2.5
1.4h	1.4135	–	–	–	4.0
6v	–	6.925	–	–	2.5
6h	–	6.925	–	–	4.0
7v	–	7.3	–	–	2.5
7h	–	7.3	–	–	4.0
10v	–	10.65	10.65	–	2.5
10h	–	10.65	10.65	–	4.0
19v	–	18.7	18.7	19.35	2.5
19h	–	18.7	18.7	19.35	4.5
24v	–	23.8	23.8	22.235	2.5
24h	–	23.8	–	–	5.0
37v	–	36.5	36.5	37.0	4.0
37h	–	36.5	36.5	37.0	7.0
89v	–	89.0	89.0	91.655	4.5
89h	–	89.0	89.0	91.655	10.0
166v	–	–	166.5	–	12.0
166h	–	–	166.5	150.0	12.0
183pm7v	–	–	183.31	–	12.0
183pm7h	–	–	–	183.31	8.0
183pm3v	–	–	183.31	–	8.0
183pm3h	–	–	–	183.31	4.0
183pm1v	–	–	–	–	4.0
183pm1h	–	–	–	183.31	2.0

Aside: what is the point of sea ice?

Add coverage

Add sensitivity

Activating atmospheric microwave imager assimilation over sea ice affected areas

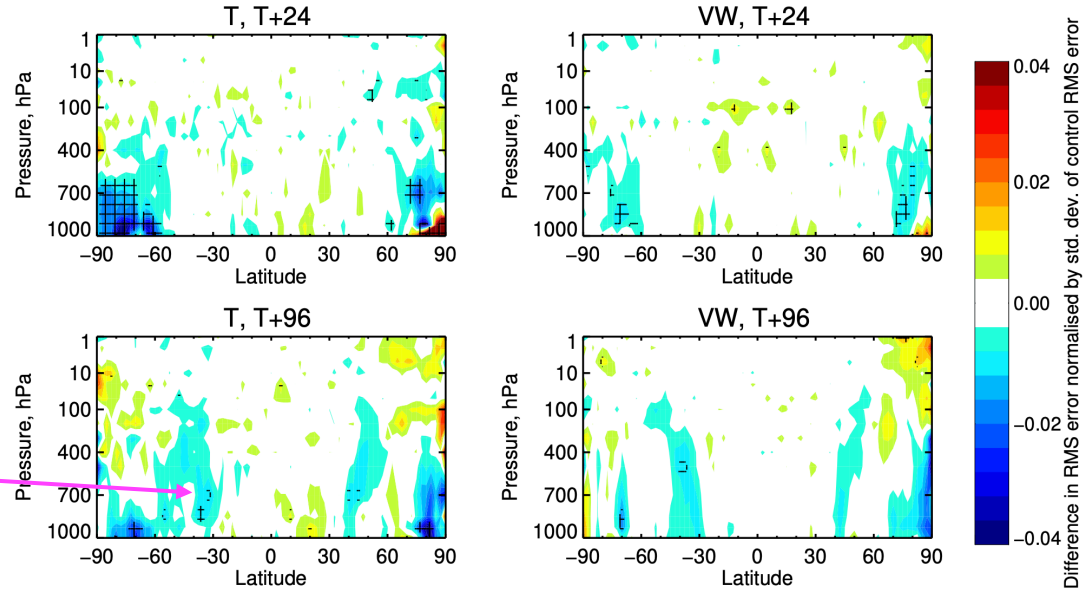


Statistically significant reduction in forecast error

Geer (2024, https://doi.org/10.1002/qj.4797)

Possibly the first demonstrated improvement in midlatitude medium-range forecasts caused by better sea ice fields

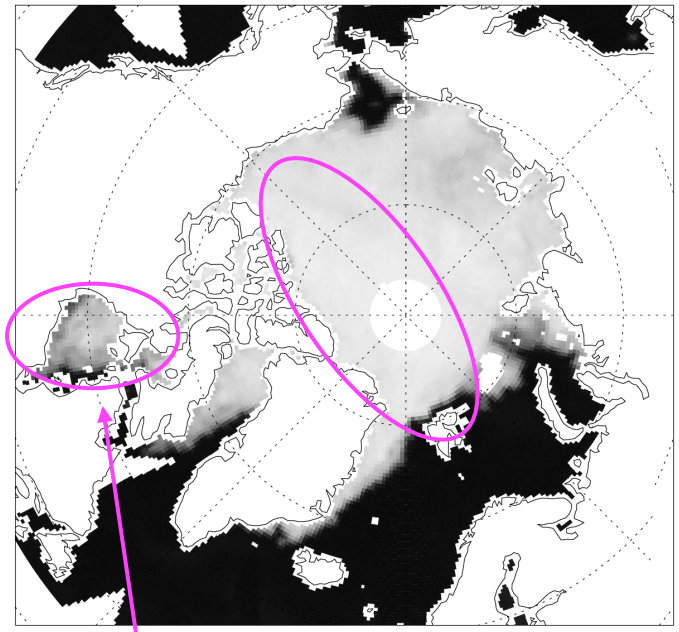
Activating coupled data assimilation of SIC into NEMO



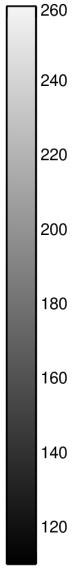
Geer, Browne, Scanlon (2026, https://doi.org/10.5194/egusphere-2026-3421)
Browne et al. (2026, https://doi.org/10.1002/qj.70266)

SIC is great, but what extra sensitivity can we get?

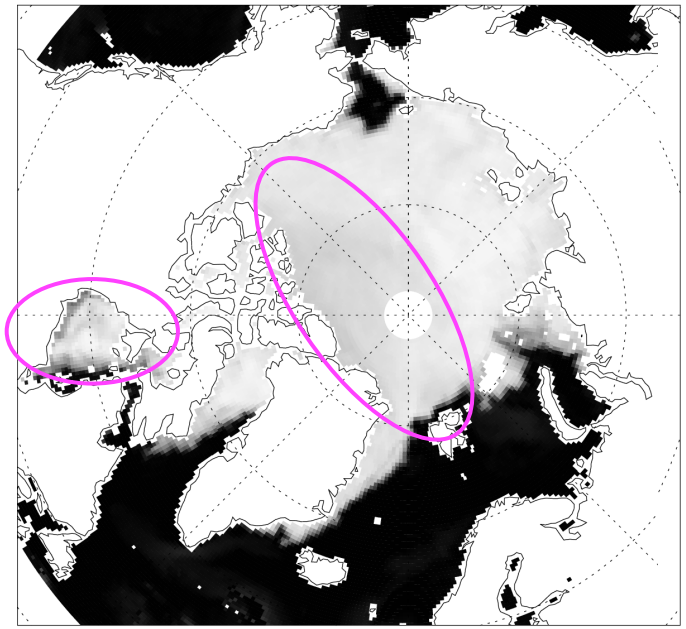
1.4 GHz v



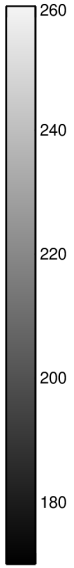
TB



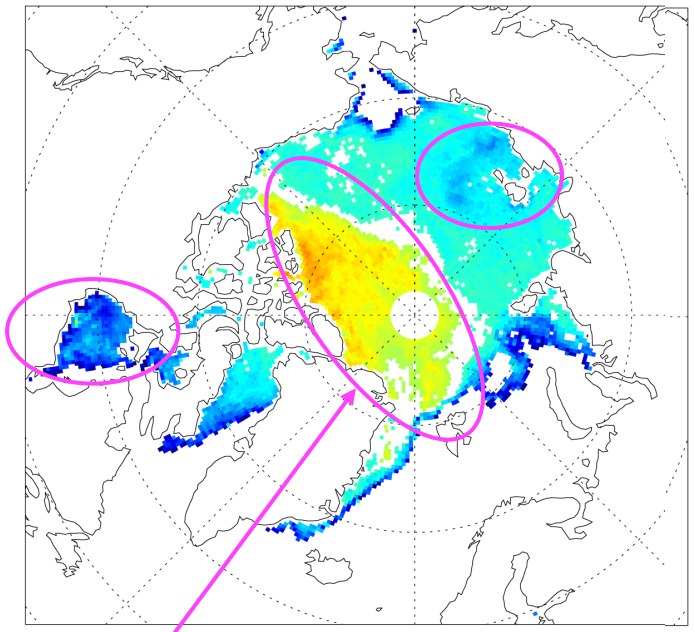
10.65 GHz v



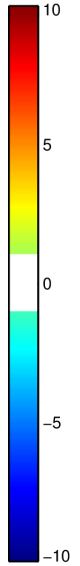
TB



6.9v minus 10.65v (sea ice only)



ΔTB



Thin new ice:

- 1.4 to 6.9 GHz sees through to the ocean beneath the ice → radiatively cold
- At 10.65 GHz sea ice is mostly opaque and emissive → radiatively warm

Thick multi year ice:

- 1.4 to 6.9 GHz sees basal ice → warm
- 10.65 GHz sees ice near surface → cold

Example: SMAP and AMSR2, 8th December 2020

To add sensitivity to more physical targets →

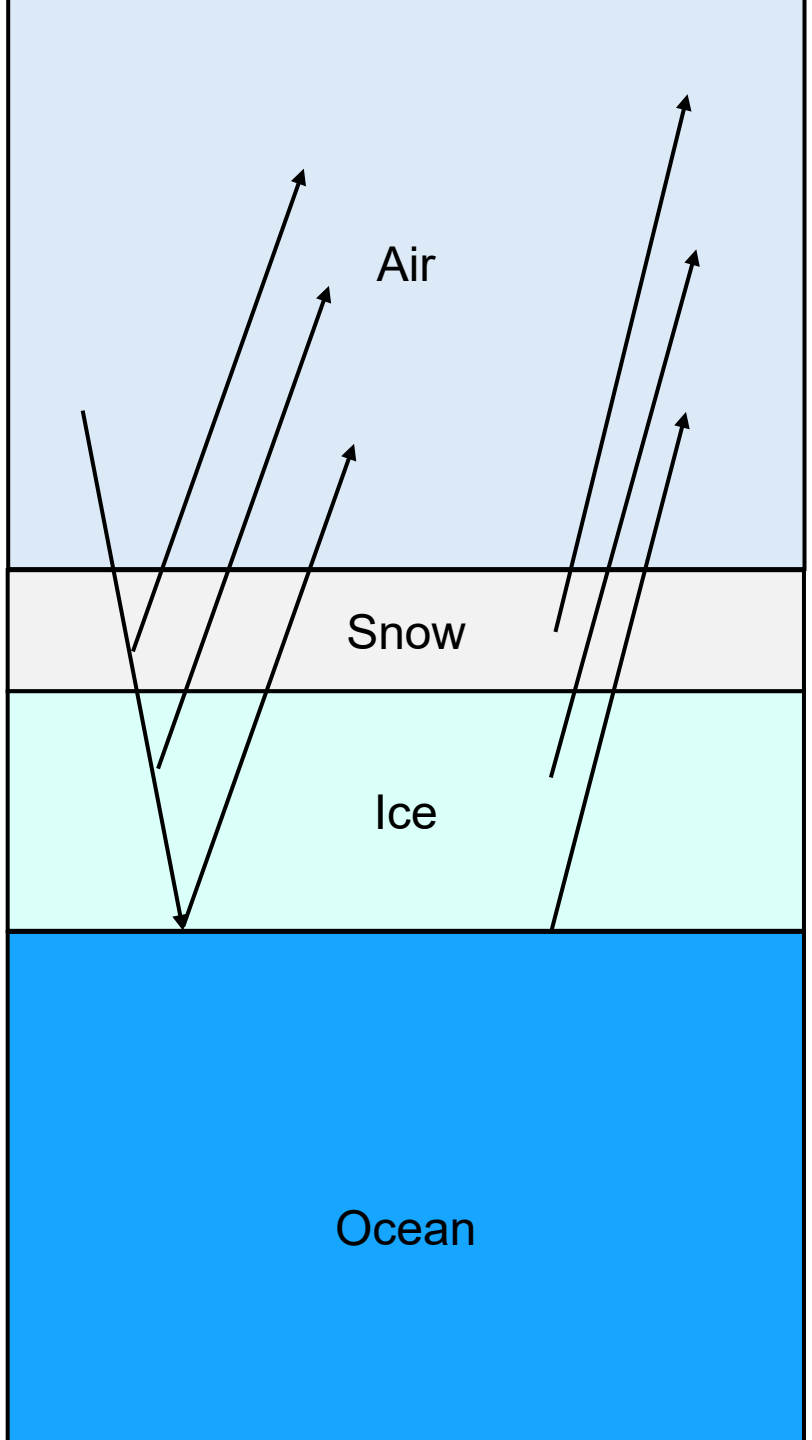
use more physics

Three empirical properties



Slab temperature (from thermodynamic model) and frequency

Slab optical properties (spectral and polarisation dependence)



Reflection and emission into the atmosphere

Two slab, two stream scattering radiative transfer model

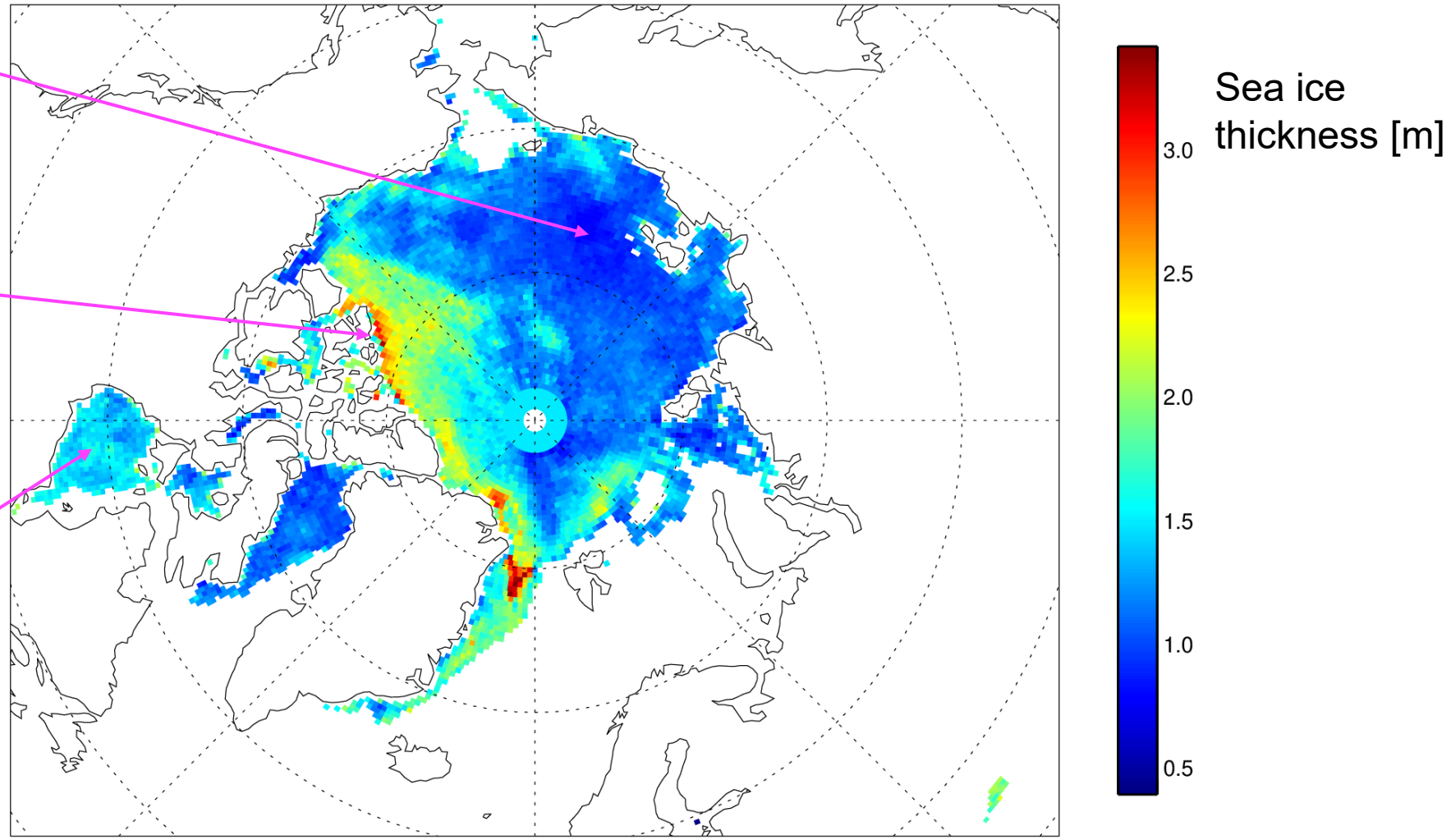
Reflection and emission at ocean boundary

Sea ice thickness retrieval from ‘v4’ more physical hybrid

Thin patch in newly formed ice

Thick multi year ice

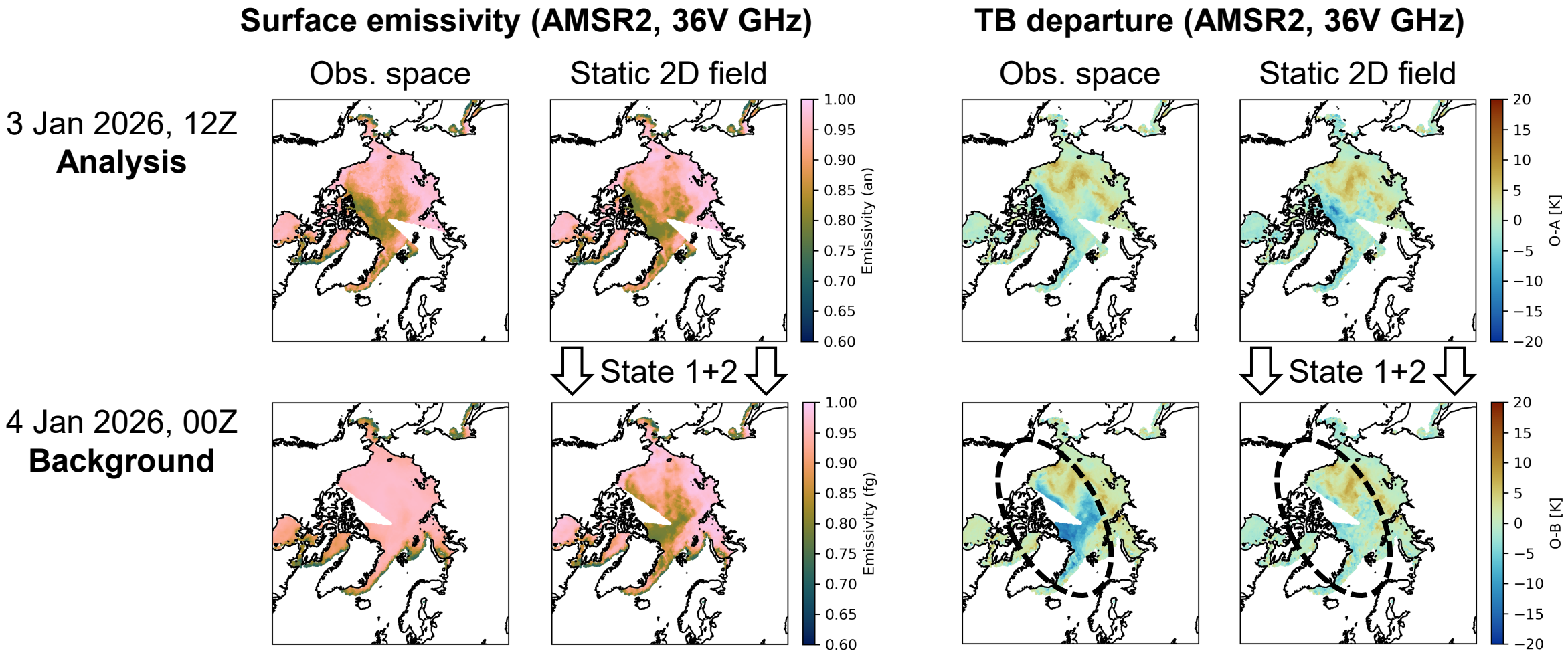
Implausibly thick 1.5 m newly formed ice?



Example: 8th December 2020. SIC > 0.5 only

Empirical state can be propagated forward in time

- 4D-Var IFS example (Nils Risse's ongoing work in the DANTEX project, using 2D-XCV)
- Memory of sea ice radiative properties improves background departure in the next cycle



Outgrowth of the sea ice work



- ESA DANTEX project
  - Prepare for assimilation of CIMR → assimilate SMAP and low-frequency AMSR channels over sea ice
    - Upgrade to v3 sea ice emissivity model (current operations is v2)
    - 2D-XCV cycling of the empirical state is necessary to assimilate SMAP
  - Develop sea ice thickness assimilation for the CRISTAL radar mission
    - When passive microwave is producing reliable thickness information, we can try that too
  - **DANTEX database** to support training wider “all-surface” hybrid forward models
    - A cycle 50r1 reanalysis from 2020 to 2025 at high resolution
    - Coupled model parameters archived in ODB - e.g. sea ice parameters, snow, upper ocean
- Coupled data assimilation
  - In the coupled IFS, extend NEMOVAR to allow direct radiance assimilation
    - Optimal extraction of the surface information content (e.g. sea ice thickness and temperature)
    - Hybrid forward models can be plugged relatively easily into NEMOVAR
- Further extend hybrid modelling for passive and active microwave:
  - Ocean surface, targeting SST
  - Land surface including snow, soil moisture and vegetation
  - Scatterometers for sea ice
  - Support cross-track microwave sounders (work at DMI)